

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-316754

(43)Date of publication of application : 16.11.1999

(51)Int.Cl.

G06F 17/00

G06F 9/44

G06F 15/18

(21)Application number : 10-123057

(71)Applicant : NEC CORP
REAL WORLD COMPUTING PARTNERSHIP

(22)Date of filing : 06.05.1998

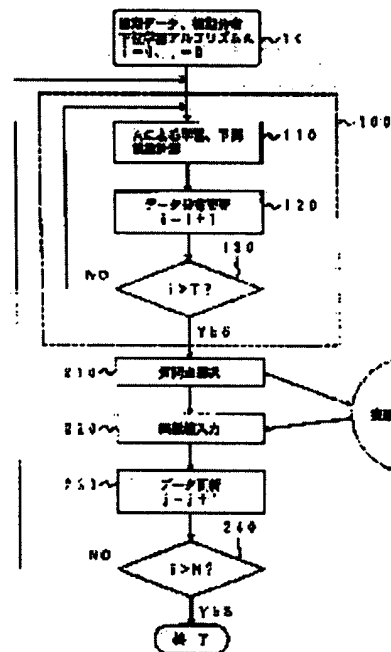
(72)Inventor : MAMIZUKA HIROSHI
ABE NAOKI

(54) EXPERIMENTAL DESIGN AND RECORDING MEDIUM RECORDING EXPERIMENTAL DESIGN PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a forming method of an experimental design, which estimates a target function with the small number of experiment times, by designating an efficient experiment procedure through the user of active learning at the time of performing an experiment where all trials becoming objects cannot practically be executed.

SOLUTION: In an input means, initial data, the initial distribution of data and a low-order learning algorithm are inputted. A boosting device 100 executes a processing, obtains a supposition, calculates a prediction value to present data from the supposition, calculates an error from the predicted value and updates the distribution of data. A means 210 uses the supposition which is finally obtained in the boosting device 100, selects a questioning point which is not contained in past data and gives the instruction of the question point as a new experiment to an experimenter. A means 220 receives a function value obtained by the experiment of the experimenter against the question point. A means 230 adds the obtained question point and the function value to the past data. When the number of question times reaches a prescribed value in a means 240, active learning is terminated. Then, the supposition obtained at that time becomes that which is to be obtained.



LEGAL STATUS

[Date of request for examination] 06.05.1998

[Date of sending the examiner's decision of rejection] 16.05.2001

[Kind of final disposal of application other than the
examiner's decision of rejection or application converted
registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection] 2001-10243[Date of requesting appeal against examiner's decision of
rejection] 15.06.2001

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-316754

(43) 公開日 平成11年(1999)11月16日

(51) IntCl.⁶

G 0 6 F 17/00

9/44

15/18

識別記号

5 5 0

5 5 0

F I

G 0 6 F 15/20

9/44

15/18

Z

5 5 0 N

5 5 0 Z

審査請求 有 請求項の数 8 O L (全 7 頁)

(21) 出願番号

特願平10-123057

(22) 出願日

平成10年(1998) 5 月 6 日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(71) 出願人 593162453

技術研究組合新情報処理開発機構

東京都千代田区東神田2-5-12 龍角散

ビル8階

(72) 発明者 馬見塚 拓

東京都港区芝五丁目7番1号 日本電気株

式会社内

(72) 発明者 安倍 直樹

東京都港区芝五丁目7番1号 日本電気株

式会社内

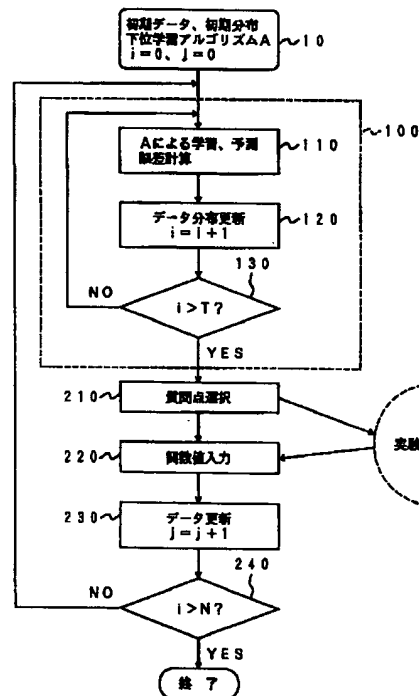
(74) 代理人 弁理士 後藤 洋介 (外1名)

(54) 【発明の名称】 実験計画法及び実験計画プログラムを記録した記録媒体

(57) 【要約】

【課題】 対象となる全試行の遂行が事実上不可能な実験を行う際に、能動学習を用いて効率的な実験手順を指定することにより、少い実験回数で目的関数の推定を可能にする実験計画の立案方法を提供する。

【解決手段】 まず、入力手段において、初期データ、そのデータの初期分布、および下位学習アルゴリズムを入力する。次に、ブースティング装置100が実行して仮説を求め、仮説から現在のデータへの予測値を計算し、予測値から誤差を計算して、データの分布を更新する。手段210では、ブースティング装置100で最終的に得られた仮説を使用して、過去のデータに含まれない質問点の選択を行い、この質問点を実験者に対する新しい実験として指示する。手段220では、上記の質問点に対して実験者の実験により得られた関数値を受け取る。手段230では、得られた質問点とその関数値を過去のデータに加える。最後に、手段240において、質問回数が一定値に達していれば、能動学習を終了する。この時点で得られる仮説が、求めたい仮説になっている。



【特許請求の範囲】

【請求項1】 入力と出力の間に正確または近似的に成り立つ関数関係について、実験によって得られる入出力のデータをもとに推定を行う実験計画法において、

次の実験の入力点の選択をする各時点で、

過去の入出力データから一定の表現形を用いて関数関係を推定する予め定められた学習アルゴリズムを下位アルゴリズムとして用いて、前記下位アルゴリズムに該時点までの入出力データを訓練データとして学習させる学習段階と、

前記学習段階の学習精度をブースティング技術により向上させるブースティング段階と、

前記ブースティング段階により複数の仮説の重み付き平均として得られる最終仮説を用いて、ランダムに生成された複数の入力候補点に対する関数値の予測を行なう関数値予測段階と、

重みの総和が最大の出力値の重み和と、重みの総和が次に大きい出力値の重み和との差が、最も小さいような入力候補点を前記入力点として選択する入力点指定段階とを行うことを特徴とする実験計画法。

【請求項2】 実験によって得られる入出力のデータを元に、入力と出力の間に成り立つ関数関係を推定する実験計画法において、

受動学習によって過去のデータから仮説を推定する仮説推定段階と、

前記仮説に基づいて次の質問点を指示する質問点指示段階と、

質問点に対する関数値を実験によって求める実験段階と、

質問点及び関数値を過去のデータに追加するデータ更新段階とを含むことを特徴とする実験計画法。

【請求項3】 請求項2記載の実験計画法において、前記仮説推定段階はブースティングを用いることを特徴とする実験計画法。

【請求項4】 請求項2及び3のいずれかに記載の実験計画法において、前記質問点指示段階は、

過去のデータにない複数の質問候補点をランダムに生成する質問候補点生成段階と、

仮説に基づいて複数の質問候補点から次の質問点を選択する質問点選択段階とを含むことを特徴とする実験計画法。

【請求項5】 入力と出力の間に正確または近似的に成り立つ関数関係について、実験によって得られる入出力のデータをもとに推定を行う実験計画プログラムを記録したコンピュータ読み取り可能な記録媒体において、

次の実験の入力点の選択をする各時点で、

過去の入出力データから一定の表現形を用いて関数関係を推定する任意の学習アルゴリズムを下位アルゴリズムとして用いて、前記下位アルゴリズムに該時点までの入出力データを訓練データとして学習させる学習処理と、

前記学習処理の学習精度をブースティング技術により向上させるブースティング処理と、

前記ブースティング処理により複数の仮説の重み付き平均として得られる最終仮説を用いて、ランダムに生成された複数の入力候補点に対する関数値の予測を行なう関数値予測処理と、

重みの総和が最大の出力値の重み和と、重みの総和が次に大きい出力値の重み和との差が、最も小さいような入力候補点を前記入力点として選択する入力点指定処理とをコンピュータに実行させることを特徴とする実験計画プログラムを記録した記録媒体。

【請求項6】 実験によって得られる入出力のデータを元に、入力と出力の間に成り立つ関数関係を推定する実験計画プログラムを記録したコンピュータ読み取り可能な記録媒体において、

受動学習によって過去のデータから仮説を推定する仮説推定処理と、

前記仮説に基づいて次の質問点を指示する質問点指示処理と、

質問点に対する関数値を実験によって求める実験処理と、

質問点及び関数値を過去のデータに追加するデータ更新処理とをコンピュータに実行させることを特徴とする実験計画プログラムを記録した記録媒体。

【請求項7】 請求項6記載の実験計画プログラムを記録した記録媒体において、前記仮説推定処理はブースティング処理を含むことを特徴とする実験計画プログラムを記録した記録媒体。

【請求項8】 請求項6及び7のいずれかに記載の実験計画プログラムを記録した記録媒体において、前記質問点指示処理は、

過去のデータにない複数の質問候補点をランダムに生成する質問候補点生成処理と、

仮説に基づいて複数の質問候補点から次の質問点を選択する質問点選択処理とをコンピュータに実行させることを特徴とする実験計画プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、分子生物学などの分野の実験を効率的に遂行する計画を立案する方法に関する。

【0002】

【従来の技術】 一般に、従来の「実験計画法」は、例えば、ある機械の性能向上、といった目的（これを特性と呼ぶ）に対し、影響を及ぼすような要因（これを因子と呼ぶ）を抽出し、因子のとり得る値（これを水準と呼ぶ）によって、特性が向上するかを検査することが目的となっている。即ち、因子のとり水準を変動させて実験することにより、因子同士の相関とその特性への影響を解析する。

【0003】しかし、従来の「実験計画法」では、実際の実験において、過去に得られた実験結果を踏まえて、次に行う実験で使用するべき因子の水準を逐一決めるわけではない。即ち、従来の「実験計画法」は、特性の追求のために、どのような実験手順を行うかという戦略を決定する方法ではない。

【0004】こうした従来の「実験計画法」において生じる問題を、一例として「免疫ペプチドの結合能の特定問題」を挙げて以下に説明する。体内のMHC細胞と呼ばれる細胞が、抗原の持つペプチドと呼ばれる小さい蛋白質のかけらと結合することによって、免疫系は抗原を認識していると考えられている。このため、「免疫ペプチドの結合能の特定問題」は、医学、特に免疫学の分野における最重要問題の一つである。例えば、ある蛋白質が、長さ10のアミノ酸配列（以下、ペプチドと呼ぶ）と結合するとし、この時アミノ酸配列に依存して結合能が変わるとする。どのようなペプチドがこの蛋白質に対して結合能が高いのかを仮に生化学実験を行って知るとすれば、長さ10のペプチドを全種類生成し、それぞれの結合能を測定する実験を行うという手順を経る必要がある。しかし、アミノ酸の種類は20なので、長さ10のペプチドには、 20^{10} の組み合わせが存在することになる。これら全ての配列に対する結合能のデータを生化学実験から得るのは到底不可能である。

【0005】つまり、より一般的に述べれば、変数 x とそれに対応するラベル（出力） y のデータから、 x と y の関係を表現する関数 $f(y=f(x))$ を知りたい時、変数 x の取り得る範囲が非常に広く、全ての x に対して $f(x)$ の値を実験的に求めることは現実的に不可能な場合が多々存在する。このような場合、従来の「実験計画法」によれば現実には実行不可能な数の実験を行わなければ解析ができない、即ち、実際には解析できなかった。

【0006】このように、従来、過去に得られたデータ (x, y) から次の実験すべき質問点を決めながら関数 f を推定していくための精度の良い方法が知られておらず、そのような方法の発明が待望されていた。

【0007】このような従来の「実験計画法」に対し、本発明での「実験計画法」は、上記「従来の実験計画法」の言葉で書けば、次のようになる。即ち、因子を入力として特性を出力とするような関数 f を推定する際に、結果を知りたい実験の因子の水準を逐一指定することにより、少ない実験回数で関数 f を推定する方法である。表現を一般化すると、入力 x とそのラベル（出力） y の関係を表す関数（仮説） $f(y=f(x))$ を効率良く推定するために行う実験の順序を計画することを指す。

【0008】本発明では、この実験計画法を能動学習により実現する。能動学習とは、例えば、1997年発行の日本の雑誌「情報処理」38巻7号558-561頁

記載の安倍と中村による解説「能動学習概要」などに記されているように、学習者が環境から選択的に情報を獲得し、学習する方法を指す。

【0009】能動学習の手法としては、多くの方法が提案されている。例えば、1995年発行の米国の雑誌「スタティスティカルサイエンス」(Statistical Science)の第10巻273-304頁記載のチャロナー(Chaloner)とベルディネリ(Verdinelli)の論文「ベイジアンエクスペリメンタルデザイン：アレビュー」(Bayesian experimental design: a Review)では、ベイズ決定理論に基づく能動学習手法が提案されている。また、1991年発行の米国の雑誌「アイトリプルイー トランザクションズ オン ニューラル ネットワークス」(IEEE Transactions on Neural Networks)の第2巻5-19頁記載のバウム(Baum)の論文「ニューラル ネット アルゴリズムズ ザット ラーン イン ポリノミアル タイム フロム イグザンプルズ アンド クエリーズ」(Neural net algorithms that learn in polynomial time from examples and queries)では、過去の学習データから境界点と推定される点を新たな質問点とする学習方法を提案している。

【0010】これらの手法と共に、データ効率性を実現するための「集団質問学習」と呼ばれる学習方法が提案されており、有力な能動学習手法の一つとされている。代表的な集団質問学習手法としては、1992年発行の国際会議の予稿集「プロシーディングス オブ ザ フィフス アニュアル エーシーエム ワークショップ オン コンピューテショナル ラーニング セオリー (Proceedings of the fifth annual ACM workshop on computational learning theory)」の287-294頁記載のセング(Seung)らの論文「クエリー バイ コミッティ」(Query by Committee)記載の方法がある。この方法では、ランダム化された複数の学習者に予測をさせ、その不一致度が最大の点を新たな質問点とし、学習を行っている。

【0011】一方、能動学習をしない学習は受動学習と呼ばれる。受動学習において、学習者は環境から情報を選択的に得ることはなく、所与の大量のデータからデータに内在する規則（仮説）を推定する。受動学習において、近年、与えられたデータからそれまでの仮説が不得意とするデータ領域を重点的に学習させるようにリサンプリングを繰り返すことによって学習精度を向上させる「ブースティング」と呼ばれる技術が注目されており、実験的にも高い性能を有することが確かめられて来ている。

る。

【0012】「ブースティング」の手法については、1995年発行の国際会議予稿集「プロシーディングス オブ ザ セカンド ヨーロピアン コンファレンス オン コンピュータショナル ラーニング セオリー (Proceedings of the second european conference on computational learning theory)」の23-37頁記載のフロイド (Freund) とシャピレ (Shapire) の論文「ア デジジョン セオレティック ジェネライゼーション オブ オンラインラーニング アンド アン アプリケーション トゥー ブースティング」(A decision-theoretic generalization of on-line learning and an application to boosting) に記載されている。更に、1996年発行の国際会議予稿集「プロシーディングス オブ ザ サード インターナショナル コンファレンス オン マシン ラーニング (Proceedings of the thirteenth international conference on machine learning)」の148-156頁記載のフロイド (Freund) とシャピレ (Shapire) の論文「エクスペリメンツ ウィズ ア ニュー ブースティング アルゴリズム」(Experiments with a new boosting algorithm) においては、受動学習における「ブースティング」技術が、そこそこの学習性能を持つ学習アルゴリズムを使用し、任意の学習精度を持つ学習アルゴリズムを構成するという、優位性を実験的に示している。

【0013】しかしながら、ブースティング技術は、受動学習でのみ使用されており、能動学習にはまだ適用されていない。即ち、能動学習においては、リサンプリングを繰り返すといった「ブースティング」技術特有の考え方はこれまで無く、この技術を能動学習の枠組で使用する方法は提案されていなかった。

【0014】

【発明が解決しようとする課題】本発明が解決しようとする課題は、対象となる全試行の遂行が事実上不可能な実験を行う際に、少ない実験回数で目的関数の推定を可能にする実験計画の立案方法、プログラム及び装置を提供することである。

【0015】

【課題を解決するための手段】本発明では、前記課題を解決するために、能動学習を用いた実験計画法を提案する。

【0016】特に、受動的な学習アルゴリズムの精度を、データのリサンプリングを繰り返すことにより向上させる手法である「ブースティング」技術を、能動学習

の有力手法の一つである集団質問学習と適切に組み合わせることにより、精度の高い能動学習を実現し、実験計画の立案に適用する。

【0017】即ち、本発明によれば、入力と出力の間に正確または近似的に成り立つ関数関係について、実験によって得られる入出力のデータをもとに推定を行う実験計画法において、次の実験の入力点の選択をする各時点で、過去の入出力データから一定の表現形を用いて関数関係を推定する予め定められた学習アルゴリズムを下位アルゴリズムとして用いて、下位アルゴリズムに該時点までの入出力データを訓練データとして学習させる学習段階と、学習段階の学習精度をブースティング技術により向上させるブースティング段階と、ブースティング段階により複数の仮説の重み付き平均として得られる最終仮説を用いて、ランダムに生成された複数の入力候補点に対する関数値の予測を行なう関数値予測段階と、重みの総和が最大の出力値の重み和と、重みの総和が次に大きい出力値の重み和との差が、最も小さいような入力候補点を前記入力点として選択する入力点指定段階とを行うことを特徴とする実験計画法が得られる。

【0018】別の言い方をすれば、実験によって得られる入出力のデータを元に、入力と出力の間に成り立つ関数関係を推定する実験計画法において、受動学習によって過去のデータから仮説を推定する仮説推定段階と、仮説に基づいて次の質問点を指示する質問点指示段階と、質問点に対する関数値を実験によって求める実験段階と、質問点及び関数値を過去のデータに追加するデータ更新段階とを含むことを特徴とする実験計画法が得られる。ここで、仮説推定段階においてブースティングを用いることが好ましい。また、質問点指示段階は、過去のデータにない複数の質問候補点をランダムに生成する質問候補点生成段階と、仮説に基づいて複数の質問候補点から次の質問点を選択する質問点選択段階とからなる。

【0019】また、本発明によれば、入力と出力の間に正確または近似的に成り立つ関数関係について、実験によって得られる入出力のデータをもとに推定を行う実験計画プログラムを記録したコンピュータ読み取り可能な記録媒体において、次の実験の入力点の選択をする各時点で、過去の入出力データから一定の表現形を用いて関数関係を推定する任意の学習アルゴリズムを下位アルゴリズムとして用いて、前記下位アルゴリズムに該時点までの入出力データを訓練データとして学習させる学習処理と、学習処理の学習精度をブースティング技術により向上させるブースティング処理と、ブースティング処理により複数の仮説の重み付き平均として得られる最終仮説を用いて、ランダムに生成された複数の入力候補点に対する関数値の予測を行なう関数値予測処理と、重みの総和が最大の出力値の重み和と、重みの総和が次に大きい出力値の重み和との差が、最も小さいような入力候補点を前記入力点として選択する入力点指定処理とをコン

ピュータに実行させることを特徴とする実験計画プログラムを記録した記録媒体が得られる。

【0020】別の言い方をすれば、実験によって得られる入出力のデータを元に、入力と出力の間に成り立つ関数関係を推定する実験計画プログラムを記録したコンピュータ読み取り可能な記録媒体において、受動学習によって過去のデータから仮説を推定する仮説推定処理と、仮説に基づいて次の質問点を指示する質問点指示処理と、質問点に対する関数値を実験によって求める実験処理と、質問点及び関数値を過去のデータに追加するデータ更新処理とをコンピュータに実行させることを特徴とする実験計画プログラムを記録した記録媒体が得られる。ここで、仮説推定処理はブースティング処理を含むことが好ましい。また、質問点指示処理は、過去のデータにない複数の質問候補点をランダムに生成する質問候補点生成処理と、仮説に基づいて複数の質問候補点から次の質問点を選択する質問点選択処理とからなる。

【0021】

【発明の実施の形態】本発明の第1の実施の形態である実験計画法アルゴリズムについて説明する。

【0022】本アルゴリズムは、受動学習を行う下位学習アルゴリズムと、これを使用して、能動学習を行う上位アルゴリズムに分かれる。

【0023】まず、ラベル付きの学習データをSとする。データ数をmとすれば、

$$S = \{ (x_1, y_1), \dots, (x_i, y_i), \dots, (x_m, y_m) \}$$

である。例えば、このデータは、前述の蛋白質とそれに結合するペプチドの例においては、 x_i がペプチド、 y_i が結合能を表す数値に相当する。

【0024】以下、下位学習アルゴリズムをAと呼ぶ。Aの役割は、所与の大量のデータ(S)から学習を行い仮説を出力することと、所与の一つの質問点 x_i に対して、学習を行った仮説を使用して、そのラベル y_i の予測値を出力することの二つである。ここで使用するAは、そこそこの学習性能を持つ学習アルゴリズムであればいずれでも良い。例えば、決定木を学習するアルゴリズムや階層型のニューラルネットワークの学習アルゴリズムなどが考えられる。ただし、以下の説明では、ラベル y_i には、 $y_i \in \{0, 1\}$ という制限を付ける。即ち、ラベルは0あるいは1の値しかとらず、また、Aは与えられた質問点に対して0あるいは1を予測値として出力する。

$$\begin{aligned} D_{t+1}(x_i) &= \frac{D_t(x_i) \cdot \beta_t}{Z} \quad \text{if } h_t(x_i) = y_i \\ &= \frac{D_t(x_i)}{Z} \quad \text{otherwise.} \end{aligned}$$

(2) (1) で得られたブースティングの最終仮説 h

【0025】次に、上位アルゴリズムについて説明する。能動学習を行う際の質問回数をN、上位アルゴリズムで使用するブースティングの際のデータのリサンプリングの回数をT、質問点を選ぶ際の候補の数をRとする。上位アルゴリズムの能動学習は、以下の(1)から(5)の手順をN回繰り返す。

【0026】(1) まず最初に、データの分布を変えながらリサンプリングを繰り返すことにより下位アルゴリズムの学習精度を増強する「ブースティング」を行う。ここで、t回目のデータの分布を D_t とする。m個の過去のデータを入力するとすれば、次の数1のように、初期分布(D_1)は一様とする。

【0027】

【数1】

$$\forall i \leq m, D_1(x_i) = \frac{1}{m}$$

「ブースティング」は、以下のIとIIをT回繰り返す。

【0028】I、分布 D_t に従って入力データからリサンプリングされたデータをもとに下位学習アルゴリズムAに学習をさせ、Aが出力した仮説と学習データから、この仮説の予測誤差を計算する。即ち、t回目の出力仮説を h_t とすれば、t回目の予測誤差 c_t を次の数2のように計算する。

【0029】

【数2】

$$c_t = \sum_{h_t(x_i) \neq y_i} D_t(x_i)$$

II、この予測誤差を使用して、Aが出力した仮説が正しく予測した点の確率値を一定倍率減らし、分布の全体を正規化することにより、次のデータの分布を定義する。即ち、予測誤差 c_t から β_t を次の数3のように計算し、 β_t を使用して、次の数4のようにt+1回目のデータの分布 D_{t+1} を計算する。ただし、数4において、Zは正規化定数である。

【0030】

【数3】

$$\beta_t = \frac{c_t}{1 - c_t}$$

【0031】

【数4】

fin(x) を、ブースティングの各回の重み log(1

／ β_t)を用いて、次の数5のように計算する。

【0032】

【数5】

$$h_{fin}(x) = \operatorname{argmax}_{y \in Y} \sum_{h_t(x)=y} \log \frac{1}{\beta_t}$$

(3) 過去のデータにない、R個の質問候補点 $C \subset X$ を

$$x_{min} = \operatorname{argmax}_{x \in C} \left| \sum_{h_t(x)=0} \log \frac{1}{\beta_t} - \sum_{h_t(x)=1} \log \frac{1}{\beta_t} \right|$$

(4) x_{min} の関数値を質問し、実験により $f(x_{min})$ を得る。

【0034】(5) 過去のデータに(4)において実験により得られた質問点の値もデータとして加える。

【0035】次に、本発明の第2の実施の形態である実験計画法について図1を参照して説明する。

【0036】本実施の形態では、まず、入力手段において、初期データ、そのデータの初期分布、および下位学習アルゴリズムを入力する。

【0037】次に、能動学習における質問の繰り返しに入るが、繰り返しの最初の部分は、ブースティング装置100で実行される。装置100においても、手段110と手段120によりデータのリサンプリングが繰り返される。まず、手段110において、下位学習アルゴリズムを使用して現在のデータからの学習が行われ、学習結果として仮説を得る。更に、手段110において、仮説から現在のデータへの予測値を計算し、予測値から誤差を計算する。次に、手段120において、手段110で得られた誤差から、データのリサンプリング分布を更新する。最後に、手段130において、リサンプリング回数が一定値に達していれば、ブースティングを終了し、そうでなければ手段110に戻り繰り返すを行う。

【0038】手段210では、ブースティング装置100で最終的に得られた仮説を使用して、過去のデータに含まれない、質問点の選択を行い、この質問点を実験者に対する新しい実験として指示する。

【0039】手段220では、実験者の実験により試された質問点の関数値を受け取る。

【0040】手段230では、得られた質問点とその関数値を過去のデータに加える。

【0041】最後に、手段240において、質問回数が

ランダムに生成し、その中から次の質問点を選択する。

質問点の選択においては、各候補点に対し、データの関数値が0である場合と1の場合それぞれのブースティングの重み和を計算し、その差を最小とする点を、次の数6のように次の質問点 $x_{min} \in C$ として選択する。

【0033】

【数6】

一定値に達していれば、能動学習を終了する。この時点で得られる仮説を最終仮説として出力する。

【0042】以上、本発明を実施の形態に基づいて説明したが、本発明はこれに限定されるものではなく、当業者の通常の知識の範囲内でその変更や改良が可能であることは勿論である。

【0043】

【発明の効果】本発明では、受動学習において、ほとんどの精度を上げることが可能な学習アルゴリズムであれば任意の精度にまで予測精度を向上させることが可能なブースティング技術を、能動学習の手法である集団質問学習と適切な方法により組み合わせることにより、精度の高い能動学習手法を実現する。

【0044】更に、場合の数が多いあるいは一回の試行の負担が大きいために、現実的にすべての試行を行うことが不可能な実験に対して、本方式を実験計画の立案に適用することにより、効率的でかつ精度の高い実験計画の実現が可能となる。

【図面の簡単な説明】

【図1】本発明の第2の実施の形態の動作を示す流れ図である。

【符号の説明】

10 入力手段

100 ブースティング実行装置

110 下位アルゴリズム学習、予測、誤差計算手段

120 データ分布更新手段

130 ブースティング実行回数判定手段

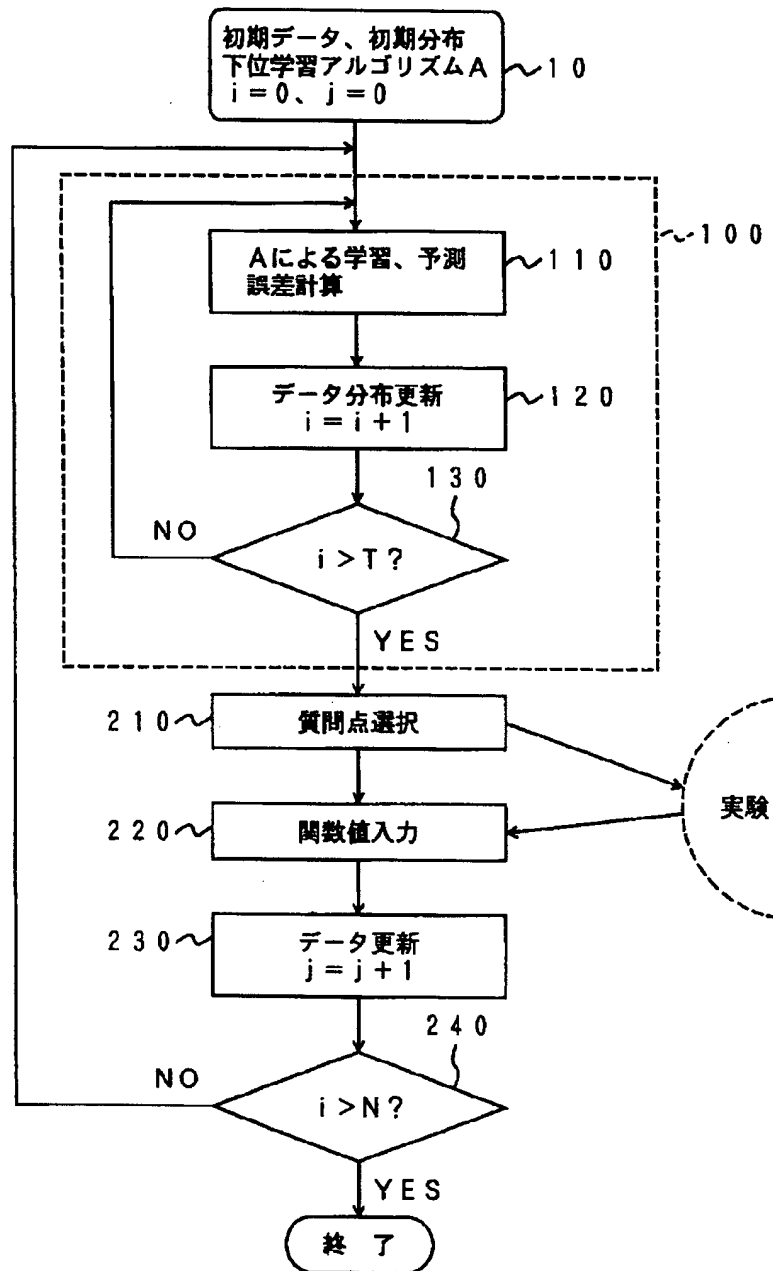
210 質問点選択手段

220 関数値入力手段

230 データ更新手段

240 能動学習実行回数判定手段

【図1】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.